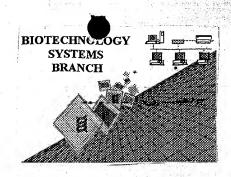# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/692,077_

Source: _OIPE_

Date Processed by STIC: _11/2/2000_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: <u>patin21help@uspto.gov</u> or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: <u>patin30help@uspto.gov</u> or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

<110> ← *insert numeric identifier and applicant names*

~~Application Project~~ *delete*

<120> ~~Title~~ : Alpha-2B-Adrenergic Receptor Polymorphisms
<130> ~~AppFileReference~~ : 13105
<140> ~~CurrentAppNumber~~ :
<141> ~~CurrentFilingDate~~ : ~~(---)~~ *delete*

<160> ← *insert numeric identifier and total number of sequences figure*

~~Sequence~~ <210> 1
<211> 1353
<212> DNA

<213> ~~OrganismName~~ : Homo sapiens
<400> ~~PreSequenceString~~ : 1

ATGGACCACC AGGACCCTA CTCCGTGCAG GCCACAGCGG CCATAGCGGC GGCCATCACC ← 60 *insert*
~~60~~TTCCTCATTC TCTTTACCAT CTTCGGCAAC GCTCTGGTCA TCCTGGCTGT GTTGACCAGC ← 120 *cumulative base totals at right margin of each line.*
~~120~~CGCTCGCTGC GCGCCCCTCA GAACCTGTTC CTGGTGTCGC TGGCCGCCGC CGACATCCTG
~~180~~GTGGCCACGC TCATCATCCC TTTCTCGCTG GCCAACGAGC TGCTGGGCTA CTGGTACTTC
~~240~~CGGCGCACGT GGTGCGAGGT GTACCTGGCG CTCGACGTGC TCTTCTGCAC CTCGTCCATC
~~300~~GTGCACCTGT GCGCCATCAG CCTGGACCGC TACTGGGCCG TGAGCCGCGC GCTGGAGTAC   *DO NOT use upper-case base letters in new Sequence Rules format. use lower-case letters for bases*
~~360~~AACTCCAAGC GCACCCCGCG CCGCATCAAG TGCATCATCC TCACTGTGTG GCTCATCGCC
~~420~~GCCGTCATCT CGCTGCCGCC CCTCATCTAC AAGGGCACC AGGGCCCCCA GCCGCGCGGG
~~480~~CGCCCCCAGT GCAAGCTCAA CCAGGAGGCC TGGTACATCC TGGCCTCCAG CATCGGATCT
~~540~~TTCTTTGCTC CTTGCCTCAT CATGATCCTT GTCTACCTGC GCATCTACCT GATCGCCAAA
~~600~~CGCAGCAACC GCAGAGGTCC CAGGGCCAAG GGGGGGCCTG GCAGGGTGA GTCCAAGCAG
~~660~~CCCCGACCCG ACCATGGTGG GGCTTTGGCC TCAGCCAAAC TGCCAGCCCT GGCCTCTGTG
~~720~~GCTTCTGCCA GAGAGGTCAA CGGACACTCG AAGTCCACTG GGGAGAAGGA GGAGGGGGAG
~~780~~ACCCCTGAAG ATACTGGGAC CCGGGCCTTG CCACCCAGTT GGGCTGCCCT TCCCAACTCA
~~840~~GGCCAGGGCC AGAAGGAGGG TGTTTGTGGG GCATCTCCAG AGGATGAAGC TGAAGAGGAG
~~900~~GAAGAGGAGG AGGAGGAGGA GGAAGAGTGT GAACCCCAGG CAGTGCCAGT GTCTCCGGCC
~~960~~TCAGCTTGCA GCCCCCCGCT GCAGCAGCCA CAGGGCTCCC GGGTGCTGGC CACCCTACGT
~~1020~~GGCCAGGTGC TCCTGGGCAG GGGCGTGGGT GCTATAGGTG GGCAGTGGTG GCGTCGAAGG
~~1080~~GCGCAGCTGA CCCGGGAGAA GCGCTTCACC TTCGTGCTGG CTGTGGTCAT TGGCGTTTTT
~~1140~~GTGCTCTGCT GGTTCCCCTT CTTCTTCAGC TACAGCCTGG GCGCCATCTG CCCGAAGCAC
~~1200~~TGCAAGGTGC CCCATGGCCT CTTCCAGTTC TTCTTCTGGA TCGGCTACTG CAACAGCTCA
~~1260~~CTGAACCCTG TTATCTACAC CATCTTCAAC CAGGACTTCC GCCGTGCCTT CCGGAGGATC
~~1320~~CTGTGCCGCC CGTGGACCCA GACGGCCTGG TGA   1353

~~<212> Type : DNA~~
~~<211> Length : 1353~~
~~SequenceName : Sequence 1~~
~~SequenceDescription :~~

delete totals

<210> 2

~~Custom Codon~~
~~1~~

~~Sequence Name : Sequence 1~~
<211> 1344
<212> DNA
~~Sequence 1~~

<213> ~~OrganismName~~ : Homo sapiens
<400> ~~PreSequenceString~~ : 2

ATGGACCACC AGGACCCCTA CTCCGTGCAG GCCACAGCGG CCATAGCGGC GGCCATCACC
~~60~~TTCCTCATTC TCTTTACCAT CTTCGGCAAC GCTCTGGTCA TCCTGGCTGT GTTGACCAGC
~~120~~CGCTCGCTGC GCGCCCCTCA GAACCTGTTC CTGGTGTCGC TGGCCGCCGC CGACATCCTG
~~180~~GTGGCCACGC TCATCATCCC TTTCTCGCTG GCCAACGAGC TGCTGGGCTA CTGGTACTTC
~~240~~CGGCGCACGT GGTGCGAGGT GTACCTGGCG CTCGACGTGC TCTTCTGCAC CTCGTCCATC
~~300~~GTGCACCTGT GCGCCATCAG CCTGGACCGC TACTGGGCCG TGAGCCGCGC GCTGGAGTAC
~~360~~AACTCCAAGC GCACCCCGCG CCGCATCAAG TGCATCATCC TCACTGTGTG GCTCATCGCC
~~420~~GCCGTCATCT CGCTGCCGCC CCTCATCTAC AAGGGCACC AGGGCCCCCA GCCGCGCGGG
~~480~~CGCCCCCAGT GCAAGCTCAA CCAGGAGGCC TGGTACATCC TGGCCTCCAG CATCGGATCT
~~540~~TTCTTTGCTC CTTGCCTCAT CATGATCCTT GTCTACCTGC GCATCTACCT GATCGCCAAA

delete

insert cumulative base totals at right margin and use lower-case letters for bases

delete (handwritten, left margin)

insert base (handwritten, right margin)
to tals ord

use
lower-case
letters.

```
 600 CGCAGCAACC GCAGAGGTCC CAGGGCCAAG GGGGGGCCTG GGCAGGGTGA GTCCAAGCAG
 660 CCCCGACCCG ACCATGGTGG GGCTTTGGCC TCAGCCAAAC TGCCAGCCCT GGCCTCTGTG
 720 GCTTCTGCCA GAGAGGTCAA CGGACACTCG AAGTCCACTG GGGAGAAGGA GGAGGGGGAG
 780 ACCCCTGAAG ATACTGGGAC CCGGGCCTTG CCACCCAGTT GGGCTGCCCT TCCCAACTCA
 840 GGCCAGGGCC AGAAGGAGGG TGTTTGTGGG GCATCTCCAG AGGATGAAGC TGAAGAGGAG
 900 GAGGAGGAGG AGGAAGAGTG TGAACCCCAG GCAGTGCCAG TGTCTCCGGC CTCAGCTTGC
 960 AGCCCCCCGC TGCAGCAGCC ACAGGGCTCC CGGGTGCTGG CCACCCTACG TGGCCAGGTG
1020 CTCCTGGGCA GGGGCGTGGG TGCTATAGGT GGGCAGTGGT GGCGTCGAAG GGCGCAGCTG
1080 ACCCGGGAGA AGCGCTTCAC CTTCGTGCTG GCTGTGGTCA TTGGCGTTTT TGTGCTCTGC
1140 TGGTTCCCCT TCTTCTTCAG CTACAGCCTG GGCGCCATCT GCCCGAAGCA CTGCAAGGTG
1200 CCCCATGGCC TCTTCCAGTT CTTCTTCTGG ATCGGCTACT GCAACAGCTC ACTGAACCCT
1260 GTTATCTACA CCATCTTCAA CCAGGACTTC CGCCGTGCCT TCCGGAGGAT CCTGTGCCGC
1320 CCGTGGACCC AGACGGCCTG GTGA                                  1344
```

~~<212> Type : DNA~~
~~<211> Length : 1344~~
~~SequenceName : Sequence 2~~
~~SequenceDescription :~~

~~Custom Codon~~
~~------------~~
~~Sequence Name : Sequence 2~~

see next page (handwritten)

<210> 7
<211> 450
<212> PRT
Sequence

<213> ~~OrganismName~~ : Homo sapiens
<400> ~~PreSequenceString~~ : 7
MDHQDPYSVQ ATAAIAAAIT FLILFTIFGN ALVILAVLTS RSLRAPQNLF LVSLAAADIL
60 VATLIIPFSL ANELLGYWYF RRTWCEVYLA LDVLFCTSSI VHLCAISLDR YWAVSRALEY
120 NSKRTPRRIK CIILTVWLIA AVISLPPLIY KGDQGPQPRG RPQCKLNQEA WYILASSIGS
180 FFAPCLIMIL VYLRIYLIAK RSNRRGPRAK GGPGQGESKQ PRPDHGGALA SAKLPALASV
240 ASAREVNGHS KSTGEKEEGE TPEDTGTRAL PPSWAALPNS GQGQKEGVCG ASPEDEAEEE
300 EEEEEEEEEC EPQAVPVSPA SACSPPLQQP QGSRVLATLR GQVLLGRGVG AIGGQWWRRR
360 AQLTREKRFT FVLAVVIGVF VLCWFPFFFS YSLGAICPKH CKVPHGLFQF FFWIGYCNSS
420 LNPVIYTIFN QDFRRAFRRI LCRPWTQTAW
~~<212> Type : PRT~~
~~<211> Length : 450~~
        ~~SequenceName : Sequence 7~~
        ~~SequenceDescription :~~

(handwritten annotations)
per Sequence Rules use three-letter amino acids and number them under every 5 amino acids.

delete

~~450~~ delete

DO NOT use TAB codes between amino acid nos. Use space characters.

a maximum of 16 amino acids per line allowed.

FYI: only three sequences shown as a sample of global errors in the entire Sequence Listing. Please consult new Sequence Rules for acceptable format, and consult sample Sequence Listing, attached.

FYI

/Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Appendix A To Subpart C to Part 1—Sample Sequence Listing

<110> Smith, John

Smith, Jane

<120> Example of a Sequence Listing

<130> 01-00001

<140> US 08/999,999

<141> 1998-02-28

<150> EP 91000000

<151> 1997-12-31

Consult

<160> 2

<170> PatentIn ver. 2.0

<210> 1

<211> 403

<212> DNA

<213> Paramecium aurelia

<220>

<221> CDS

<222> 341..394

<300>

<301> Doe, Richard

<302> Isolation and Characterization of a Gene Encoding a

Protease from Paramecium sp.

<303> Journal of Fictional Genes

<304> 1

<305> 4

<306> 1 - 7

<307> 1988-06-20

<400> 1

```
ctactctact ctactctcat ctactatctt ctttggatct ctgagtctgc ctgagtggta    60

ctcttgagtc ctggagatct ctcctctcac atgtgatcgt cgagactgac cgatagatcg   120

ctgactgact ctgagatagt cgagcccgta cgagacccgt cgagggtgac agagagtggg   180

cgcgtgcgcg cagagcgccg cgccggtgcg cgcgcgagtg cgcggtgggc cgcgcgaggg   240

ctttcgcggc agcggcggcg ctttccggcg cgcgcccgtc cgcccctaga cctgagaggt   300

cttctcttcc ctcctcttca ctagagaggt ctatatatac atg gtt tca atg ttc    355
                                               Met Val Ser Met Phe
```

agc ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgtttgctc   403

Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val

                  10   15

<210> 2

<211> 18

<212> PRT

<213> Paramecium aurelia

<400> 2

Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu

1   5   10   15

Phe Val

identifiers and their accompanying information as shown in the following
table. The numeric identifier shall be used only in the "Sequence
Listing." The order and presentation of the items of information in the
"Sequence Listing" shall conform to the arrangement given below. Each
item of information shall begin on a new line and shall begin with the
numeric identifier enclosed in angle brackets as shown. The submission
of those items of information designated with an "M" is mandatory. The
submission of those items of information designated with an "O" is
optional. Numeric identifiers <110> through <170> shall only be set
forth at the beginning of the "Sequence Listing." The following table
illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials.	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:#:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	O
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	O
<302>	Title		O
<303>	Journal		O
<304>	Volume		O
<305>	Issue		O
<306>	Pages		O
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	O
<308>	Database Accession Number	Accession number assigned by database including database name	O
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	O
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	O

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	O
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	O
<313>	Relevant Residues	FROM (position) TO (position)	O
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	M

5. Section 1.824 is revised to read as follows:

   1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by   1.821(e) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a)(1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;